

ARGUMENTS/REMARKS

Claims 1 through 11 and 13 through 23 are pending in the present application. Claims 21-23 are new. Claims 1, 18 and 20 are independent. Claims 1, 18 and 21 have been amended by the present amendment. Claim 12 has been cancelled without prejudice.

Applicants wish to thank the Examiner for the many courtesies extended in scheduling and conducting the telephonic interview. Present on the call were Examiner Abebe, and Mme. Wingood and Gat-Falik.

During the interview, a proposed version of claim 1 was discussed with reference to Gorin. The proposed claim 1 was amended relative to the then-currently-pending claim 1 to detail the segmentation step to include a parameterization step, an anchoring step, and a modeling and classification step.

In the Office Action, claims 1-20 were rejected under § 102(e) as being anticipated by U.S. Pat. 7,295,970 to Gorin, et al. ("Gorin").

Present claim 1 is directed to a speaker segmentation method for associating one segment for each of two sides of a summed audio interaction, with one of the at least two sides of the interaction using additional information. Claim 1 further provides a receiving step, a segmentation step and a scoring step.

Present claim 1 further provides that the segmentation step include a parameterization step for transforming the speech signal into a set of feature vectors and dividing the set into non-overlapping segments; an anchoring step for locating an anchor segment for each side of the interaction; and a modeling and classification step for associating at least one second segment with each side of the interaction.

Thus, claim 1 now provides for dividing the audio signal into segments, and then selects **two specific segments** out of the segment set, as a starting point for the speaker segmentation, wherein one of the two specific segments represents one participant of the interaction, and the other represents the second participant.

In Gorin, however, the interaction is divided into segments based on speaker change, not divided based upon non-overlapping segments in which only a single speaker is speaking, as presently claimed. Further, in Gorin no particular segment is specifically selected out of all segments. Gorin then continues to perform clustering of all the segments. This can be seen in Fig. 4 of Gorin, wherein all segments are involved in the clustering all along the process.

Thus, Gorin does not disclose an anchoring step for locating an anchor segment for each side of the interaction.

Claim 1 further provides that the speaker segmentation method uses additional information. Gorin does not disclose using any information beyond the recorder interaction for the segmentation.

Thus, Gorin does not disclose claim 1. Reconsideration and withdrawal of the 102 rejection of claim 1 are respectfully requested. Claims 18 and 20, addressed below are directed to apparatus claims.

Claims 2 through 11 and 13 through 17 depend from independent claim 1 and, for at least the reason of such dependence, are also patentable over the cited art of record. The dependent claims contain additional features absent from the prior art of record. For example:

Dependent claim 2 provides additional information sources, such as computer telephony integration, data obtained by processing, and the like, none of which is disclosed by Gorin. Gorin assuming that in the first segment it is the agent who speaks uses a hard-coded rule, and not data or selected from any of the mentioned sources. Additionally, Gorin's assumption may prove wrong, for example in situations in which the customer starts speaking immediately, alone or together with the agent. Such wrong initial assumption can severely harm the results.

Dependent claims 4 through 8 require that the scoring step use discriminative information for discriminating the at least two sides of the interaction. In contrast,

Gorin uses a hard-coded rule, indicating that the agent is the first to speak, rather than relevant data.

Dependent claims 7 through 11 and 13 provide additional information sources to be used during the scoring step, the anchoring step or the modeling and classification step. Dependent claims 8 and 14 provides additional information sources, now of which are disclose in the Gorin.

Claim 10 requires repeating the segmentation step and the scoring step if the segmentation score is below a threshold. For example, Gorin teaches stopping the clustering when the difference between consecutive iterations is small, i.e., the clustering converges. (Col. 8, lines 25 through 32). However, Gorin does not comparing the segmentation results to an external threshold, such as the thresholds mentioned in claim 11.

Claim 15 requires a preprocessing step for enhancing the quality of the interaction, which is not disclosed by Gorin.

Further, claim 16 requires eliminating non-speech segments. In contrast, Gorin at Col. 9 mentions non-speech segments as speaker changes, but does not mention eliminating them.

The same arguments as for claim 1 are also applicable towards claim 18.

Independent claim 18 is directed to a speaker segmentation apparatus for associating an at least one segment for each of at least two speakers participating in an at least one audio interaction. The apparatus of claim 18 provides for, in pertinent part, a parameterization component for transforming the speech signal into a set of feature vectors and *dividing* the set into non-overlapping segments; an anchoring component for locating an anchoring segment for each of the at least two sides of the interaction; and a modeling and classification component for associating at least one second segment with each side of the interaction.

Therefore present claim 18 is allowable at least for the reasons set above for claim 1.

Claims 19 and 23 depend from claim 18, and for at least the reason of such dependence, are also patentable over the cited art.

Present claim 20 is allowable at least for the reasons set above for claim 1. Independent claim 20 is directed to a quality management apparatus for interaction-rich environments. The apparatus provides for, in pertinent part, a parameterization component for transforming the speech signal into a set of feature vectors and *dividing* the set into non-overlapping segments; an anchoring component for locating an anchoring segment for each of the at least two sides of the interaction; and a modeling and classification component for associating at least one second segment with each side of the interaction.

New dependent claim 21 depends from independent claim 1 for provides an embodiment for the anchoring step, in which the anchor segment for one speaker is selected as a homogenous segment, a statistical model is constructed from the first anchor segment, and then the anchor segment for the second speaker is selected such that its statistical model is as different as possible from the statistical model of the first anchor segment.

New dependent claim 22 depends from claim 21 and provides that the homogeneous segment be selected by spotting a predetermined sequence.

The claimed selection of the anchor segments provides an efficient starting point, in which there is high probability that each of the two segments belongs to a single speaker, and that each of them belongs to a different speaker. Gorin does not disclose selecting two specific segments, let alone selecting the segments in this specific manner.

Applicants submit that the application is now in condition for allowance. Passage of the claims to allowance is respectfully requested.

Respectfully submitted,

March 16, 2009
Date

Charles N.J. Ruggiero
Charles N.J. Ruggiero
Reg. No. 28,468
Attorney for the Applicants
Ohlandt, Greeley, Ruggiero & Perle, L.L.P.
One Landmark Square, 10th Floor
Stamford, CT 06901-2682
Tel: 203-327-4500
Fax: 203-327-6401